

10/660,639

11-14.03

DIALOG(R) File 351:Derwent WPI
(c) 2003 Thomson Derwent. All rts. reserv.

014604454 **Image available**

WPI Acc No: 2002-425158/200245

XRPX Acc No: N02-334296

Image processing apparatus executes drawing process in such a way that a character or character train recognized by analyzer is output with input font size

Patent Assignee: CANON KK (CANO); HINO Y (HINO-I)

Inventor: HINO Y

Number of Countries: 002 Number of Patents: 003

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
US 20020036788	A1	20020328	US 2001950020	A	20010912	200245 B
JP 2002091946	A	20020329	JP 2000277245	A	20000912	200245
JP 2002189582	A	20020705	JP 2001274082	A	20010910	200259

Priority Applications (No Type Date): JP 2000277245 A 20000912; JP 2000275952 A 20000912

Patent Details:

Patent No	Kind	Lan Pg	Main IPC	Filing Notes
US 20020036788	A1	50	G06K-001/00	
JP 2002091946	A	23	G06F-017/21	
JP 2002189582	A	34	G06F-003/12	

Abstract (Basic): US 20020036788 A1

NOVELTY - An analyzer analyzes document data and recognizes the font size information contained in the document data. An input unit is provided to input standard font size allotted to a physical page. A data drawer executes a drawing process in such a way that a character or character train recognized by the analyzer is output with the input font stage.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- (a) Image processing method;
- (b) Computer readable memory storing image processing program;
- (c) Server apparatus

USE - Image processing apparatus for copier, printer, inkjet printer, facsimile, etc.

ADVANTAGE - Achieves appropriate formatting of document utilizing information contained in structured document and/or information relating to physical page to which structured document is to be allocated. The image processing apparatus is capable of realizing page layout of high freedom, thereby selecting output according to taste of various generation from aged person to children.

DESCRIPTION OF DRAWING(S) - The figure shows the flowchart explaining data processing sequence in server apparatus.

pp; 50 DwgNo 13/42

Title Terms: IMAGE; PROCESS; APPARATUS; EXECUTE; DRAW; PROCESS; WAY; CHARACTER; CHARACTER; TRAIN; RECOGNISE; ANALYSE; OUTPUT; INPUT; FONT; SIZE

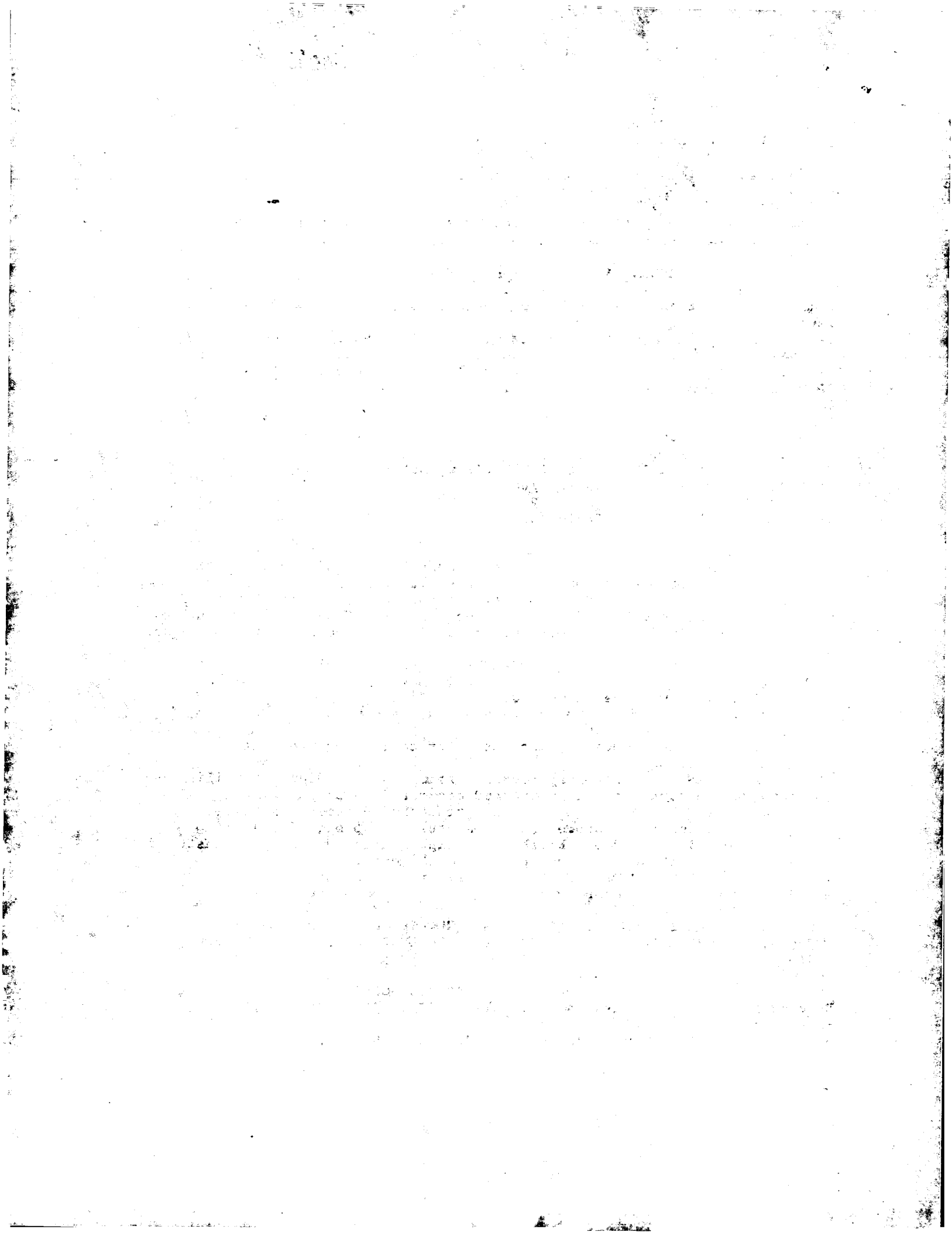
Derwent Class: P75; T01; W02

International Patent Class (Main): G06F-003/12; G06F-017/21; G06K-001/00

International Patent Class (Additional): B41J-021/00; H04N-001/393

File Segment: EPI; EngPI

Manual Codes (EPI/S-X): T01-J10C3; T01-J11A; T01-S03; W02-J03A2B



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-91946

(P2002-91946A)

(43) 公開日 平成14年3月29日 (2002.3.29)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード* (参考)
G 0 6 F 17/21	5 3 0	G 0 6 F 17/21	5 3 0 T 2 C 0 8 7
	5 4 0		5 4 0 2 C 1 8 7
B 4 1 J 21/00		B 4 1 J 21/00	Z 5 B 0 0 9

審査請求 未請求 請求項の数16 O L (全 23 頁)

(21) 出願番号 特願2000-277245(P2000-277245)

(22) 出願日 平成12年9月12日 (2000.9.12)

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 日野 康弘

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(74) 代理人 100076428

弁理士 大塚 康徳 (外2名)

Fターム(参考) 2C087 AA03 AA09 AA18 AB06 AC08

BA03 BA05 BA06 BA14 BD06

CA03

2C187 AC07 AD03 AE06

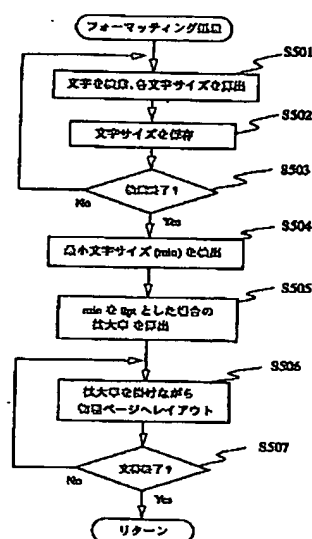
5B009 NB11 NC06 NE03

(54) 【発明の名称】 画像処理装置及びその方法、及び画像処理システム

(57) 【要約】

【課題】 構造化記述言語によって記述された文書データの物理ページへのレイアウト処理は、固定倍率によるものであり、ユーザの好みに応じた適切なレイアウト結果が得られるとは限らなかった。

【解決手段】 構造化記述言語で記述された文書データを物理ページにレイアウトする際に、前記文書データから特定データとして最小文字サイズを検出し (S504)、該検出された特定データをレイアウトするための変倍率を算出し (S505)、前記文書データに対して、物理ページへのレイアウトが可能となるように前記変倍率に基づく変換を施す (S506)。



【特許請求の範囲】

【請求項1】 構造化記述言語で記述された文書データを物理ページにレイアウトする画像処理装置であって、前記文書データから特定データを検出し、該検出された特定データをレイアウトするための変倍率を算出する算出手段と、

前記文書データに対して、物理ページへのレイアウトが可能となるように前記変倍率に基づく変換を施す変換手段と、を有することを特徴とする画像処理装置。

【請求項2】 前記特定データは、前記文書データ中における最小文字サイズであることを特徴とする請求項1記載の画像処理装置。

【請求項3】 前記特定データは、前記文書データ中における最大オブジェクト幅であることを特徴とする請求項1記載の画像処理装置。

【請求項4】 前記特定データは、前記文書データ中における最頻出文字サイズであることを特徴とする請求項1記載の画像処理装置。

【請求項5】 前記算出手段は、前記特定データのサイズを所定サイズとするための変倍率を算出することを特徴とする請求項1記載の画像処理装置。

【請求項6】 更に、ユーザ指示に基づいてレイアウト方式を選択する選択手段を備え、前記算出手段は、前記選択手段における選択結果に基づいて前記変倍率の算出方法を決定することを特徴とする請求項1記載の画像処理装置。

【請求項7】 前記算出手段は、前記選択手段における選択結果に基づいて、前記文書データから検出すべき特定データの種類の決定することを特徴とする請求項6記載の画像処理装置。

【請求項8】 前記算出手段は、前記選択手段における選択結果に基づいて、前記特定データとして前記文書データ中における最小文字サイズ、最大オブジェクト幅、最頻出文字サイズのうちいずれかを決定することを特徴とする請求項7記載の画像処理装置。

【請求項9】 前記構造化記述言語はHTMLであることを特徴とする請求項1記載の画像処理装置。

【請求項10】 前記構造化記述言語はXMLであることを特徴とする請求項1記載の画像処理装置。

【請求項11】 更に、前記変換手段によって変換された文書データを描画する描画手段を有することを特徴とする請求項1記載の画像処理装置。

【請求項12】 更に、前記描画手段によって描画されたデータを記録媒体上に可視像化する画像形成手段を有することを特徴とする請求項1記載の画像処理装置。

【請求項13】 構造化記述言語で記述された文書データを物理ページにレイアウトする画像処理方法であって、

前記文書データから特定データを検出し、該検出された特定データをレイアウトするための変倍率を算出する算

出ステップと、

前記文書データに対して、物理ページへのレイアウトが可能となるように前記変倍率に基づく変換を施す変換ステップと、を有することを特徴とする画像処理方法。

【請求項14】 画像処理装置と画像形成装置とを接続し、構造化記述言語で記述された文書データを物理ページにレイアウトして画像形成する画像処理システムであって、

前記文書データから特定データを検出し、該検出された特定データをレイアウトするための変倍率を算出する算出手段と、

前記文書データに対して、物理ページへのレイアウトが可能となるように前記変倍率に基づく変換を施す変換手段と、を有することを特徴とする画像処理システム。

【請求項15】 構造化記述言語で記述された文書データを物理ページにレイアウトする画像処理を実現するプログラムであって、

前記文書データから特定データを検出し、該検出された特定データをレイアウトするための変倍率を算出する算出ステップのコードと、

前記文書データに対して、物理ページへのレイアウトが可能となるように前記変倍率に基づく変換を施す変換ステップのコードと、を有することを特徴とするプログラム。

【請求項16】 構造化記述言語で記述された文書データを物理ページにレイアウトする画像処理プログラムを記録した記録媒体であって、該プログラムは、

前記文書データから特定データを検出し、該検出された特定データをレイアウトするための変倍率を算出する算出ステップのコードと、

前記文書データに対して、物理ページへのレイアウトが可能となるように前記変倍率に基づく変換を施す変換ステップのコードと、を有することを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、構造化記述言語で記述された文書データを物理ページにレイアウトする画像処理装置及びその方法、及び画像処理システムに関する。

【0002】

【従来の技術】パーソナルコンピュータ及びインターネットの急速な普及により、業種を問わず文書の電子化が促進されつつある。一般に、電子化文書のデータ形式は文書を編集するアプリケーションに依存するため、電子化された文書を閲覧するためには該当文書のデータ形式をサポートするアプリケーションを用意する必要があった。

【0003】そこで、HTML(Hyper Text Markup Language)やXML(eXtensible Markup Language)といった、特定

のアプリケーションに依存しない構造化記述言語によって記述された文書が普及しつつある。一般に構造化記述言語は、ブラウザによってディスプレイ上に表示することを前提に設計されており、表示する画面の幅や高さを任意に変えたり、画面をスクロールさせたりすることが可能であるため、「ページ」という概念は必要ない。

【0004】しかしながら、構造化記述言語をページ記述言語として利用する場合には、ページへの割り付けを行う必要がある。そこで、近年ページレイアウト可能な構造化記述言語が開発されつつあり、出版用の組版ルール等を用いることによって、ページ概念の無い構造化記述言語をページレイアウト可能な構造化記述言語に変換することも可能となった。このように、ページ概念の無い構造化記述言語をページレイアウト可能な構造化記述言語に変換する工程は、「フォーマット」ing」と称される。

【0005】図9に、ページ概念のない構造化記述言語で記述された文書のブラウザによる表示例を示す。一般的なブラウザによれば同図に示すように、表示画面の大きさを変えたり、スクロールバーを使って画面を縦横スクロールさせることができる。

【0006】これに対し、図10にフォーマット後の構造化記述言語で記述された文書の表示例を示す。同図によれば、フォーマット後の文書はページ単位に表示または印刷することができる。

【0007】

【発明が解決しようとする課題】しかしながら、上記従来のフォーマットingにおいては、構造化記述言語によって記述された文書データの物理ページへの割り付けは、固定倍率によるものであった。従って、ページの内容によっては見えにくくなってしまったり、ユーザ好みのレイアウトが得られるとは限らないといった問題が発生していた。

【0008】本発明は上述した問題を解決するために成されたものであり、構造化記述言語によって記述された文書データを物理ページへレイアウトする際に、自由度の高いレイアウトを実現する画像処理装置及びその方法、及び画像処理システムを提供することを目的とする。

【0009】

【課題を解決するための手段】上記目的を達成するための一手段として、本発明の画像処理装置は以下の構成を備える。

【0010】即ち、構造化記述言語で記述された文書データを物理ページにレイアウトする画像処理装置であって、前記文書データから特定データを検出し、該検出された特定データをレイアウトするための変倍率を算出する算出手段と、前記文書データに対して、物理ページへのレイアウトが可能となるように前記変倍率に基づく変換を施す変換手段と、を有することを特徴とする。

【0011】例えば、前記特定データは、前記文書データ中における最小文字サイズであることを特徴とする。

【0012】例えば、前記特定データは、前記文書データ中における最大オブジェクト幅であることを特徴とする。

【0013】例えば、前記特定データは、前記文書データ中における最頻出文字サイズであることを特徴とする。

【0014】例えば、前記算出手段は、前記特定データのサイズを所定サイズとするための変倍率を算出することを特徴とする。

【0015】更に、ユーザ指示に基づいてレイアウト方式を選択する選択手段を備え、前記算出手段は、前記選択手段における選択結果に基づいて前記変倍率の算出方法を決定することを特徴とする。

【0016】例えば、前記算出手段は、前記選択手段における選択結果に基づいて、前記特定データとして前記文書データ中における最小文字サイズ、最大オブジェクト幅、最頻出文字サイズのうちいずれかを決定することを特徴とする。

【0017】例えば、前記構造化記述言語はHTMLであることを特徴とする。

【0018】例えば、前記構造化記述言語はXMLであることを特徴とする。

【0019】更に、前記変換手段によって変換された文書データを描画する描画手段を有することを特徴とする。

【0020】更に、前記描画手段によって描画されたデータを記録媒体上に可視像化する画像形成手段を有することを特徴とする。

【0021】

【発明の実施の形態】以下、本発明に係る一実施形態について、図面を参照して詳細に説明する。

【0022】＜第1実施形態＞まず、本実施形態を適用するに好適なレーザービームプリンタ（以下、単に「プリンタ」と記述する）の構成について、図1を参照して説明する。尚、本実施形態が適用されるプリンタはレーザービーム方式に限られるものではなく、他のプリント方式のプリンタでも良いことは言うまでもない。

【0023】図1は、本実施形態が適用されるプリンタの内部構造を示す断面図であり、不図示のデータ源からの文字パターンの登録や定型書式（フォームデータ）などの登録が行える。同図において、1000はプリンタ本体であり、外部に接続されているホストコンピュータから供給される文字情報（文字コード）やフォーム情報あるいはマクロ命令などを入力して記憶するとともに、それらの情報に従って対応する文字パターンやフォームパターンなどを作成し、記録媒体である記録紙上に像を形成する。1012は操作のためのスイッチおよびLED表示器などが配されている操作パネル、1001はプ

リント1000全体の制御およびホストコンピュータから供給される文字情報などを解析するプリンタ制御ユニットである。この制御ユニット1001は、主に文字情報を対応する文字パターンのビデオ信号に変換してレーザドライバ1002に出力する。レーザドライバ1002は半導体レーザ1003を駆動するための回路であり、入力されたビデオ信号に応じて半導体レーザ1003から発射されるレーザ光1004をオンオフ切り替える。

【0024】レーザ1004は回転多面鏡1005で左右方向に振られ静電ドラム1006上を走査する。これにより、静電ドラム1006上には文字パターンの静電潜像が形成される。この潜像は、静電ドラム1006周囲の現像ユニット1007により現像された後、記録紙に転送される。この記録紙にはカットシートを用い、カットシート記録紙はプリンタ1000に装着した用紙カセット1008に収納され、給紙ローラ1009および搬送ローラ1010と1011とにより装置内に取り込まれて、静電ドラム1006に供給される。

【0025】図2は、本実施形態におけるプリンタ制御システムの構成を説明するブロック図である。ここでは、図1に示したプリンタ1000を制御する例について説明する。尚、本発明の機能が実行されるのであれば、単体の機器であっても、複数の機器からなるシステムであっても、LAN等のネットワークを介して処理が行われるシステムであっても、本発明を適用できることは言うまでもない。

【0026】図2において、3000はホストコンピュータであり、ROM内3のプログラム用ROMに記憶された文書処理プログラム等に基づいて文書データの取り出しやデータ変換を実行するCPU1を備え、システムバス4に接続される各デバイスをCPU1が総括的に制御する。

【0027】ROM3内のプログラム用ROMには、CPU1の制御プログラム等が記憶され、同じくフォント用ROMには上記データ変換処理の際に使用するフォントデータ等が記憶され、同じくデータ用ROMは上記データ変換処理等を行う際に使用する各種データが記憶される。

【0028】2はRAMであり、CPU1の主メモリやワークエリア等として機能する。5はキーボードコントローラ(KBC)であり、キーボード9や不図示のポインティングデバイスからのキー入力を制御する。6はCRTコントローラ(CRTC)であり、CRTディスプレイ(CRT)10の表示を制御する。尚、本実施形態においてキーボード9やCRTディスプレイ10は必ずしも必要ではないが、ホストコンピュータのメンテナンスや動作状況の確認のために、通常装備されている。

【0029】7はメモリコントローラ(MC)であり、ブートプログラム、種々のアプリケーション、フォント

データ、ユーザファイル、編集ファイル等を記憶するハードディスク(HD)、フロッピー(登録商標)ディスク(FD)等の外部メモリ11とのアクセスを制御する。8はネットワークコントローラ(NTC)であり、所定の双方向性インタフェース(インタフェース)21を介してプリンタ1000に接続され、プリンタ1000との通信制御処理を実行する。

【0030】尚、CPU1は、例えばメモリコントローラ7を制御することによって外部メモリ11に記憶された文書データを取り出したり、ネットワークコントローラ8を制御することによって文書データを外部へ転送することができる。

【0031】プリンタ1000において、12はプリンタCPUであり、ROM13内のプログラム用ROMに記憶された制御プログラム或いは外部メモリ14に記憶された制御プログラム等に基づいて、システムバス15に接続される各種のデバイスとのアクセスを総括的に制御し、印刷部インタフェース16を介して接続される印刷部(プリンタエンジン)17に、出力情報としての画像信号を出力する。

【0032】ROM13内のプログラムROMには、後述するフローチャートで示されるようなCPU12の制御プログラムを記憶しても良い。同じくフォント用ROMには上記出力情報を生成する際に使用するフォントデータ等を記憶し、同じくデータ用ROMは、プリンタ1000がハードディスク等の外部メモリ14を備えない場合には、ホストコンピュータ3000上で利用される情報等を記憶している。

【0033】CPU12は、入力部18を介してホストコンピュータ3000との通信処理を行うことによって、プリンタ1000内の情報等をホストコンピュータ3000に通知することが可能である。19はCPU12の主メモリ、ワークエリア等として機能するRAMであり、図示しない増設ポートに接続されるオプションRAMによってメモリ容量を拡張することができる。尚、RAM19は出力情報展開領域、環境データ格納領域、NVRAM等に用いられる。

【0034】上述したハードディスク(HD)、ICカード等の外部メモリ14は、メモリコントローラ(MC)20によりアクセスが制御される。外部メモリ14はオプションとして接続され、文書データ、フォントデータ、フォームデータ等を記憶する。18は操作パネルであり、プリンタ1000に対する各種操作のためのスイッチおよびLED表示器等が配されている。

【0035】尚、本実施形態における外部メモリ14は1個に限らず、少なくとも1個以上備え、内蔵フォントに加えてオプションフォントカード、言語系の異なるプリンタ制御言語を解釈するプログラムを格納した外部メモリを複数接続しても良い。更に不図示のNVRAMを有し、操作パネル1012からのプリンタモード設定情

報を記憶するようにしても良い。

【0036】図3は、本実施形態におけるプリンタ1000の機能構成を示すブロック図である。同図によればプリンタ1000は、大きく分けてフォーマット制御部1100、プリンタインターフェイス1200、出力制御部1300、プリンタエンジン部1400によって構成されている。

【0037】プリンタインターフェイス1200は、フォーマット制御部1100と外部とにおけるデータの入出力を制御するための手段である。

【0038】フォーマット制御部1100は、プロトコル制御部1101、文書データ解析部1102、データ描画部1103、ページメモリ1104、物理レイアウト処理部1105により構成されている。一般的にフォーマット制御部1100内のこれらの機能構成は、CPUやROM、RAM等を備えたコンピュータシステムによって実現される。

【0039】プロトコル制御部1101は、ネットワークプロトコルを解析・送信することによって外部との通信を行なう手段であり、例えばプロトコルとしてHTTP (Hyper Text Transfer Protocol) を使用した場合、URL (Uniform Resource Locator) によって指示された文書の取得や、Webサーバへの情報の送信を行なう。

【0040】文書データ解析部1102は、構造化記述言語で記述された文書データを解析し、より処理しやすい形式の中間コードに変換する。文書データ解析部1102において生成された中間コードは、データ描画部1103に渡されてビットマップデータに展開された後、ページメモリ1104に逐次描画される。

【0041】物理レイアウト処理部1105においては、構造化記述言語によって記述された文書データの物理ページへの割り付け、即ちフォーマット処理を行なう。具体的には、文書データ解析部1102において解析された文書データがHTMLのようなページ概念がない文書データであった場合に、物理ページへの割り付けが行われたデータを作成する。

【0042】出力制御部1300は、ページメモリ1104の内容をビデオ信号に変換し、プリンタエンジン部1400への画像転送を行なう。プリンタエンジン部1400は、受け取ったビデオ信号を記録紙上に永久可視画像形成するための印刷機構部である。

【0043】次に、上述した構成からなる本実施形態の印刷システムにおける印刷制御手順を、図4乃至図6に示すフローチャートを参照して説明する。

【0044】図4は、プリンタ1000の動作開始から終了までのメイン処理を示すフローチャートである。まずステップS401で、プリンタインターフェイス1200を介してネットワークからのデータを受け取る。次にステップS402でプロトコルの解析を行なった後、ステップS403においてフォーマット処理、即

ち物理ページへの割り付けを行ない、ステップS404に進んで描画処理を行なう。その後、ステップS405で文書データが終了したか否かを判断し、終了であれば印刷動作を終了する。一方、文書データ終了でなければ、ステップS401に戻って上記処理を繰り返す。

【0045】図5は、図4のステップS403に示したフォーマット処理、即ち物理レイアウト処理部1105における物理ページへのレイアウト処理の詳細を示すフローチャートである。

【0046】まず、ステップS501で文書データを先頭から検索し、文字データを検出した時点でその文字サイズを算出し、ステップS502で該算出した文字サイズを保存する。そして、ステップS503で全ての文書データの検索が終了したか否かを判定し、終了するまでステップS501～S502の処理を繰り返すことにより、文書中に使用されている文字サイズのリストが完成する。

【0047】ステップS503で全ての文書データの検索が終了した場合にはステップS504に進み、保存された文字サイズのリストから、最小の文字サイズminを検出する。そしてステップS505において、検出された最小文字サイズminを8ポイントと見立てた場合の拡大率を算出する。この拡大率が物理レイアウトの基準となるため、ステップS506ではこの拡大率を用いて、文書の先頭から順次、物理ページへの割り付けを行なう。

【0048】そしてステップS507において、全ての文書データに対して物理ページへの割り付けが終了したか否かを判断し、終了するまでステップS506の割り付け処理を繰り返す。

【0049】ここで、上記図5に示すフォーマット処理の具体例を示す。ステップS501～S503によって、文書中に使用されている文字サイズとして例えば、6、8、10ポイントがリストされた場合、ステップS504では最小文字サイズminとして6ポイントが検出される。するとステップS505において、これを8ポイントとするための拡大率が $8/6$ 倍として算出される。従って、リストされた他の8、10ポイントの文字サイズはそれぞれ、 $8 \times (8/6)$ 、 $10 \times (8/6)$ ポイントで物理ページに割り付けられる。尚、文字のみでなくテーブルやボーダーライン等、他の図形(オブジェクト)についてもこの拡大率(上記例では $8/6$)を掛け合わせることによって、物理ページへ割り付けられる。

【0050】図6は、図4のステップS404に示した描画処理の詳細を示すフローチャートである。尚、この描画処理は即ちデータ描画部1103における描画処理であり、実際に印刷を行う処理である。

【0051】まずステップS601において、文書データ解析部1102でページ終了タグが検出されたか否か

を判別する。ページ終了タグが検出されなければステップS602に進み、次に解析したタグが文字印字または図形描画等、ページメモリ1104への展開処理を必要とするタグであるか否かを判別する。否であればステップS605に進み、属性設定や印字位置制御等、解析したタグに従った処理を直ちに実行する。

【0052】一方、印字・描画等のタグであればステップS603に進み、ビットマップへの展開処理が容易に行える形式の中間コードを生成する。そしてステップS604において、データ描画部1103でこの中間コードを受けて、ページメモリ1104へのビットマップ展開処理を行う。展開処理終了後は図4のステップS402に戻り、文書データのプロトコル解析処理を繰り返す。

【0053】一方、ステップS601においてページ終了タグが検出された場合はステップS606に進み、出力制御部1300においてページメモリ1104の内容をプリンタエンジン部1400に対するビデオ信号に変換し、画像転送出力を行う。そしてステップS607においてプリンタエンジン部1400では、受け取ったビデオ信号に基づいて記録紙上に永久可視画像を形成し、印刷を行う。そしてステップS608において印刷が行われた記録紙が排紙されると、1ページの印刷制御処理が終了する。

【0054】ここで、本実施形態におけるフォーマット処理結果の具体例を示す。図11はHTMLによって記述された入力文書データ例であり、この文書データは、本実施形態のフォーマット処理が施されることによって、図12に示すような、ページレイアウト可能な構成に変換される。また、図12に示す文書データを実際に印刷した結果を、図15に示す。

【0055】この例によれば、図11に示される最小文字("table 1")のサイズは4ポイントであるから、拡大率は $8/4=2$ 倍として算出される。従って、図11において"<H1>"のタグで示されるタイトル文字("Sample")のサイズが14ポイント、その他の文字("This document...web browser.")のサイズが規定値である6ポイントであるとする、図12によれば、タイトルの文字サイズが28ポイント、その他の文字サイズが12ポイント、そして最小文字サイズが8ポイントとして変換されていることが分かる。また、テーブルオブジェクトについてもこの拡大率(2倍)が適用されるため、テーブルが2分割でページ内に割りつけられていることが分かる。

【0056】以上説明した様に本実施形態によれば、構造化記述言語をフォーマット処理する際に、その物理レイアウトの基準値を最小文字サイズに基づいて設定することによって、ページ内への適切な割り付けが可能となる。

【0057】尚、本実施形態における拡大率の算出は、

構造化記述言語内において検出された最小文字サイズminを8ポイントに拡大する場合について説明したが、本発明はこの例に限定されず、他のサイズに拡大するように拡大率を算出する場合でも、同様に適用される。

【0058】また、算出された拡大率に基づいて文字サイズを決定する際に、単に拡大率を乗じるだけでなく、更に適切なポイント数(システムにおいて通常使用されるポイント数)へ切り上げ等の最適化処理を施しても良い。

【0059】<第2実施形態>以下、本発明に係る第2実施形態について説明する。

【0060】上述した第1実施形態においては、物理レイアウトの基準値として最小文字サイズを参照する例について説明した。第2実施形態では、該基準値をオブジェクトの最大幅に基づいて設定することを特徴とする。ここで、オブジェクトとは文字以外の描画図形を指し、HTML文書ではテーブルやイメージ等がこれにあたる。

【0061】図7は、第2実施形態におけるフォーマット処理を示すフローチャートである。尚、第2実施形態におけるシステム構成、及び印刷動作のメイン処理、描画処理については、上述した第1実施形態と同様であるため、説明を省略する。

【0062】まずステップS701において文書データを先頭から検索し、文字以外のオブジェクトを検索する。例えばHTMLであれば、"<TABLE>"や""というタグを検索すれば良い。

【0063】次にステップS702において、検出したオブジェクトの幅を算出する。例えばオブジェクトがテーブルであれば、該テーブルに指定された幅情報、もしくはテーブル内に記述される文字サイズに基づいて幅が算出できる。また、イメージであれば、データフォーマットのヘッダに記述される画像情報に基づいて幅が算出できる。

【0064】そしてステップS703において、全ての文書データの検索が終了したか否かを判定し、終了するまでステップS701～S702の処理を繰り返すことにより、文書中に使用されているオブジェクトの幅のリストが完成する。

【0065】ステップS703で全ての文書データの検索が終了した場合にはステップS704に進み、保存されたオブジェクト幅のリストから、最大のオブジェクト幅maxを検出する。そしてステップS705において、検出された最大オブジェクト幅maxをページ幅と比較する。ここでページ幅とは、記録紙内において物理的に印字可能となる主捜査方向の長さを指す。

【0066】最大オブジェクト幅maxがページ幅以下であればステップS707に進み、縮小率を1に設定する。一方、最大オブジェクト幅maxがページ幅よりも大きければステップS706に進み、最大オブジェクト幅maxをページ幅とみなした場合の縮小率を算出する。こ

のようにして求められた縮小率が物理レイアウトの基準となるため、ステップS708ではこの縮小率を用いて、文書の先頭から順次、物理ページへの割り付けを行なう。

【0067】そしてステップS709において、全ての文書データに対して物理ページへの割り付けが終了したか否かを判断し、終了するまでステップS708の割り付け処理を繰り返す。

【0068】尚、文字以外のオブジェクトを備えない、即ち全てが文字によって構成されている文書データについては、ステップS701においてオブジェクトが検出されない。このような場合、図7のフローチャートには特に示していないが、縮小率を1に設定して、物理ページへの割り付け処理を行なう。そのために例えば、ステップS701でオブジェクトが検出されなかった場合、オブジェクト最大幅maxを最小値（例えば0）に設定するステップを設ければ良い。

【0069】ここで、上記図7に示すフォーマット処理の具体例を示す。ステップS701～S703によって、文書中に使用されているオブジェクト幅として例えば、2000、5000ドットがリストされた場合、ステップS704では最大オブジェクトサイズmaxとして5000ドットが検出される。一方、記録紙がA4サイズ、プリンタ解像度が600dpiであれば、該記録紙を縦に搬送する場合のページ幅は4720ドットである。するとステップS705において最大オブジェクトサイズmax（5000）がページ幅（4720）よりも大きいため、ステップS706で縮小率が4720/5000として算出される。尚、オブジェクトのみでなく文字についてもこの縮小率を掛け合わせるによって、物理ページへ割り付けられる。

【0070】ここで、第2実施形態におけるフォーマット処理結果の具体例を示す。上述した第1実施形態で図11に示したHTMLの入力文書データについて、第2実施形態のフォーマット処理を施すことによって、図13に示すような変換結果が得られる。図13に示す文書データを実際に印刷した結果は、図16に示すようになる。図13及び図16によれば、最大幅を有するオブジェクトであるテーブルがページ内に適切に割りつけられ、文字サイズが第1実施形態に示した図15よりも小さくなっていることが分かる。

【0071】以上説明した様に第2実施形態によれば、構造化記述言語をフォーマット処理する際に、その物理レイアウトの基準値を最大オブジェクト幅に基づいて設定することによって、ページ内への適切な割り付けが可能となる。

【0072】＜第3実施形態＞以下、本発明に係る第3実施形態について説明する。

【0073】第3実施形態においては、物理レイアウトの基準値を最も使用頻度の高い文字サイズに基づいて設

定することを特徴とする。

【0074】図8は、第3実施形態におけるフォーマット処理を示すフローチャートである。尚、第3実施形態におけるシステム構成、及び印刷動作のメイン処理、描画処理については、上述した第1実施形態と同様であるため、説明を省略する。

【0075】まずステップS801において文書データを先頭から検索し、文字データを検出した時点でその文字サイズを算出し、ステップS802で文字サイズ毎の文字数をカウントして保存する。そして、ステップS803で全ての文書データの検索が終了したか否かを判定し、終了するまでステップS801～S802の処理を繰り返すことにより、文書中に使用されている文字サイズ及びその文字数のリストが完成する。

【0076】ステップS803で全ての文書データの検索が終了した場合にはステップS804に進み、保存された文字サイズのリストから、最も文字数の多い、即ち最頻出の文字サイズfreqを検出する。そしてステップS805において、検出された最頻出文字サイズfreqを10ポイントと見立てた場合の拡大率を算出する。この拡大率が物理レイアウトの基準となるため、ステップS806ではこの拡大率を用いて、文書の先頭から順次、物理ページへの割り付けを行なう。

【0077】そしてステップS807において、全ての文書データに対して物理ページへの割り付けが終了したか否かを判断し、終了するまでステップS806の割り付け処理を繰り返す。

【0078】ここで、上記図8に示すフォーマット処理の具体例を示す。ステップS801～S803によって、例えば文書中に使用されている文字サイズとして8ポイントが10文字、10ポイントが400文字、20ポイントが8文字リストされた場合、ステップS804では最頻出文字サイズfreqとして10ポイントが検出される。するとステップS805において、これを10ポイントとするための拡大率が10/10倍（＝1倍）として算出される。従って、リストされた他の8、20ポイントの文字サイズもそれぞれ、8、20ポイントで物理ページに割り付けられる。尚、文字のみでなくテーブルやボーダーライン等、他の図形（オブジェクト）についてもこの拡大率（上記例では1倍）を掛け合わせるによって、物理ページへ割り付けられる。

【0079】ここで、第3実施形態におけるフォーマット処理結果の具体例を示す。上述した第1実施形態で図11に示したHTMLの入力文書データについて、第3実施形態のフォーマット処理を施すことによって、図14に示すような変換結果が得られる。図14に示す文書データを実際に印刷した結果は、図17に示すようになる。

【0080】この例によれば、図11に示される最頻出文字サイズ("This document...webbrowser.")は6ボイ

ントであるから、拡大率は10/6倍として算出される。従って、図11におけるタイトル文字("Sample")のサイズが14ポイント、最小文字("table 1")のサイズが4ポイントであるから、図14によれば、最頻出文字サイズが10ポイント、タイトルの文字サイズが $14 \times 10 / 6 = 23.33$ (23) ポイント、最小文字サイズが $4 \times 10 / 6 = 6.66$ (7) ポイントとして変換されていることが分かる。また、最大幅を有するテーブルオブジェクトについてもこの拡大率(10/6倍)が適用され、かろうじてページ内に割りつけられていることが分かる。

【0081】以上説明した様に第3実施形態によれば、構造化記述言語をフォーマットする際に、その物理レイアウトの基準値を最頻出文字サイズに基づいて設定することによって、ページ内への適切な割り付けが可能となる。

【0082】尚、本実施形態における拡大率の算出も第1実施形態と同様に、検出された最頻出文字サイズfreqを10ポイントに拡大する例に限定されず、他のサイズへの拡大を想定しても同様に適用される。また、文字サイズの変換の際にも、拡大率を乗じるのみならず、適切なポイント数への切り上げ等の最適化処理を施しても良い。

【0083】<第4実施形態>以下、本発明に係る第4実施形態について説明する。

【0084】第4実施形態においては、上述した第1乃至第3実施形態に示したフォーマット処理のいずれかを、任意に選択可能とすることを特徴とする。即ち、第4実施形態におけるシステム構成、及び印刷動作のメイン処理、描画処理については、上述した第1実施形態と同様であるが、フォーマット処理における物理レイアウトの基準値を、最小文字サイズ、最大オブジェクト幅、又は最頻出文字サイズのいずれかに基づいて設定することを特徴とする。

【0085】具体的には、プリンタ1000の操作パネル1012において、ユーザがフォーマットの基準値として上記3種類のいずれかを選択する、即ち、第1乃至第3実施形態に対応するレイアウト方法のいずれかを選択できるように構成すれば良い。

【0086】また、選択した基準値に基づくフォーマット結果を、ホストコンピュータに接続されたCRT10においてページ単位にプレビューできるようにすれば、更に操作性が向上する。

【0087】以上説明した様に第4実施形態によれば、ユーザの所望するようなフォーマットが可能となる。例えば、図11に示すHTMLで記述された文書データについて、図15乃至図17のいずれかの出力形態が選択可能となる。

【0088】<他の実施形態>上述した第1及び第3実施形態においては、文字サイズをポイント値として算出

する例について説明したが、例えば文字サイズがその他の表現(例えば大/中/小)によって指定されている文書であっても、予め各サイズ(大/中/小)の比率を保持しておくことにより、本発明を適用することが可能である。

【0089】また、上述した各実施形態においては、フォーマット対象となる文書データとしてHTMLによって記述された文書を例として説明したが、本発明はこの例に限定されるものではなく、構造化記述言語であれば、XMLやSGMLといった他の言語であっても良いことは言うまでもない。

【0090】なお、本発明は、複数の機器(例えばホストコンピュータ、インタフェイス機器、リーダ、プリンタなど)から構成されるシステムに適用しても、一つの機器からなる装置(例えば、複写機、ファクシミリ装置など)に適用してもよい。

【0091】また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体(または記録媒体)を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ(またはCPUやMPU)が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム(OS)などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0092】さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0093】

【発明の効果】以上説明したように本発明によれば、構造化記述言語によって記述された文書データを物理ページレイアウトする際に、自由度の高いレイアウトが実現できる。

【図面の簡単な説明】

【図1】本発明に係る一実施形態におけるプリンタの側断面図である。

【図2】本実施形態における印刷システムの構成を示す

ブロック図である。

【図3】本実施形態におけるプリンタの機能構成を示すブロック図である。

【図4】本実施形態における印刷処理を示すフローチャートである。

【図5】本実施形態におけるフォーマット処理を示すフローチャートである。

【図6】本実施形態における描画処理を示すフローチャートである。

【図7】第2実施形態におけるフォーマット処理を示すフローチャートである。

【図8】第3実施形態におけるフォーマット処理を示すフローチャートである。

【図9】ページ概念のない構造化記述言語による文書データの表示例を示す図である。

【図10】図9に示す文書データのページレイアウト後の表示例を示す図である。

【図11】HTMLで記述された文書データ例を示す図である。

【図12】本実施形態におけるフォーマット後の文書データ例を示す図である。

【図13】第2実施形態におけるフォーマット後の

の文書データ例を示す図である。

【図14】第3実施形態におけるフォーマット後の文書データ例を示す図である。

【図15】本実施形態におけるフォーマット後の印刷例を示す図である。

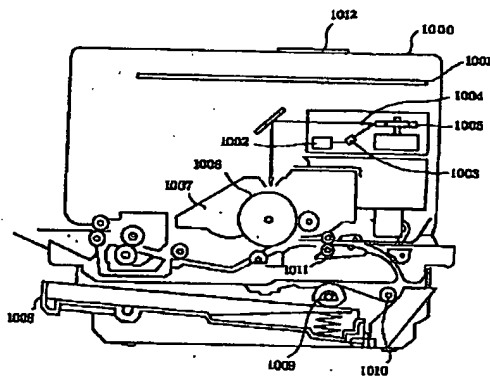
【図16】第2実施形態におけるフォーマット後の印刷例を示す図である。

【図17】第3実施形態におけるフォーマット後の印刷例を示す図である。

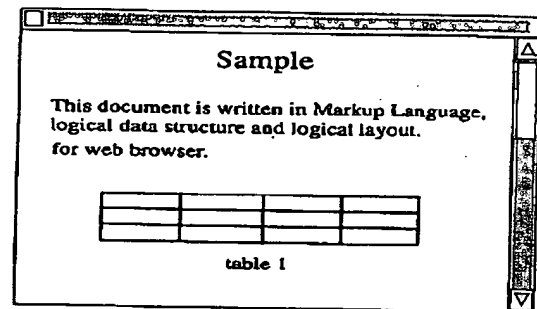
【符号の説明】

1000 プリンタ
1012 操作パネル
1100 フォーマット制御部
1101 プロトコル制御部
1102 文書データ解析部
1103 データ描画部
1104 ページメモリ
1105 物理レイアウト処理部
1200 プリンタインタフェース
1300 出力制御部
1400 プリンタエンジン部

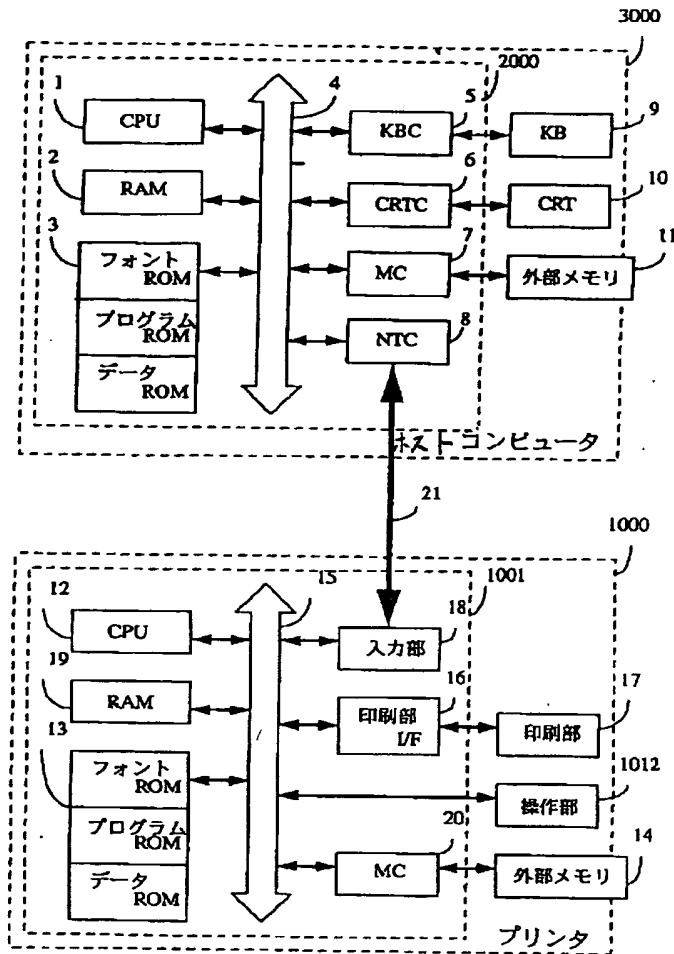
【図1】



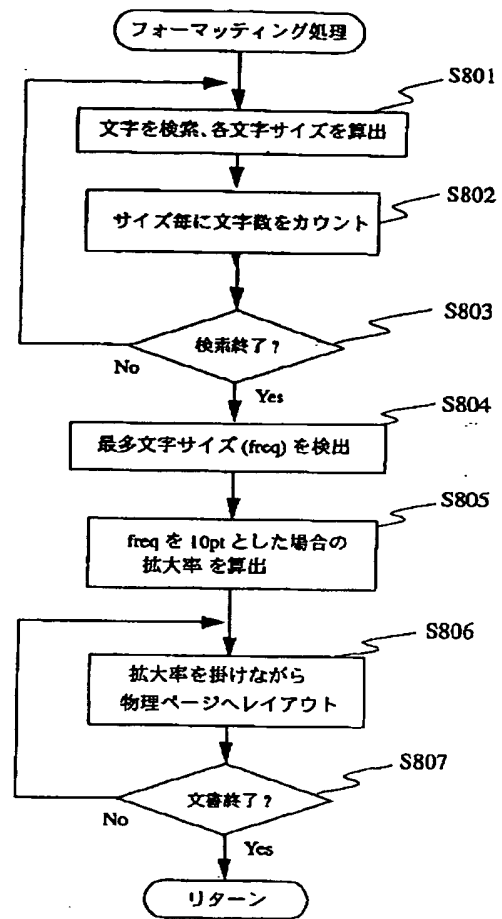
【図9】



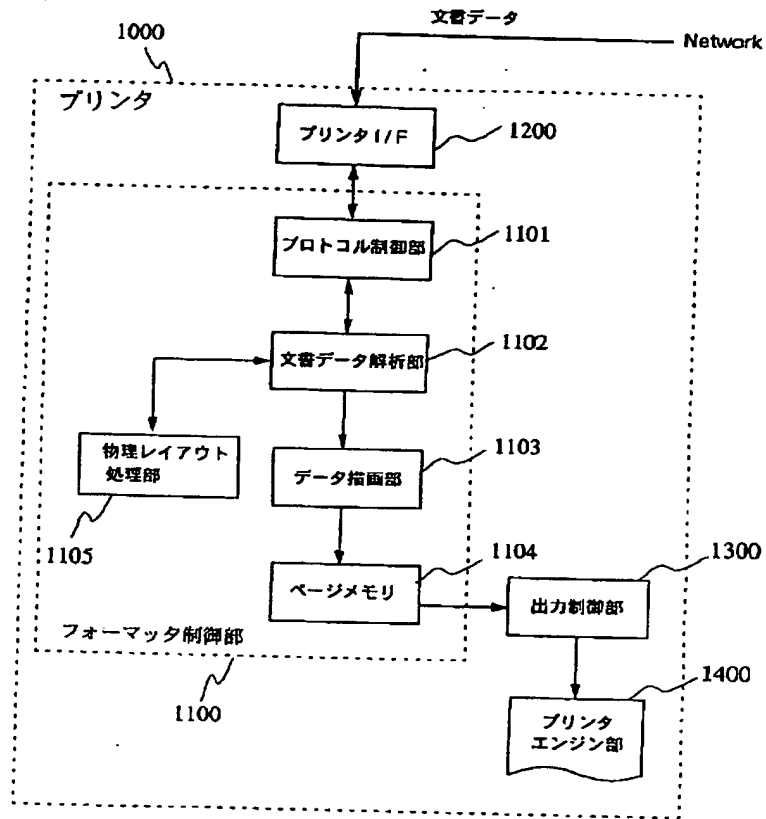
【図2】



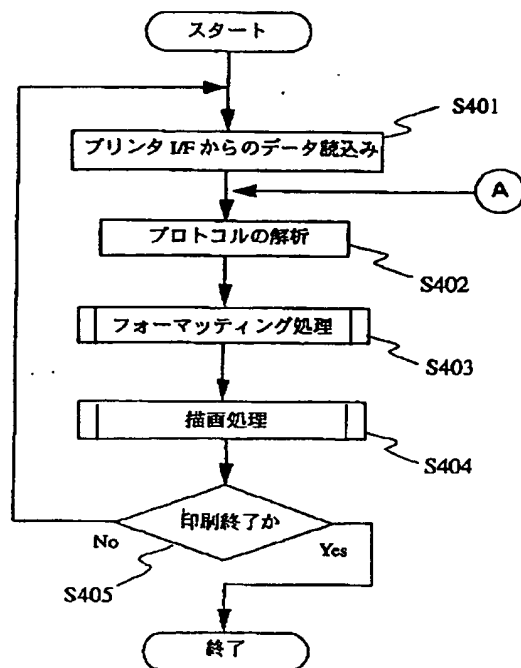
【図8】



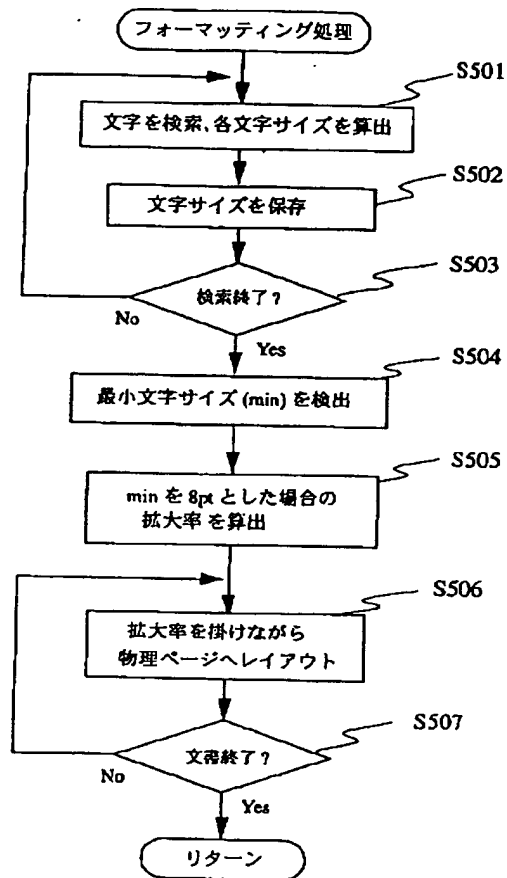
【図3】



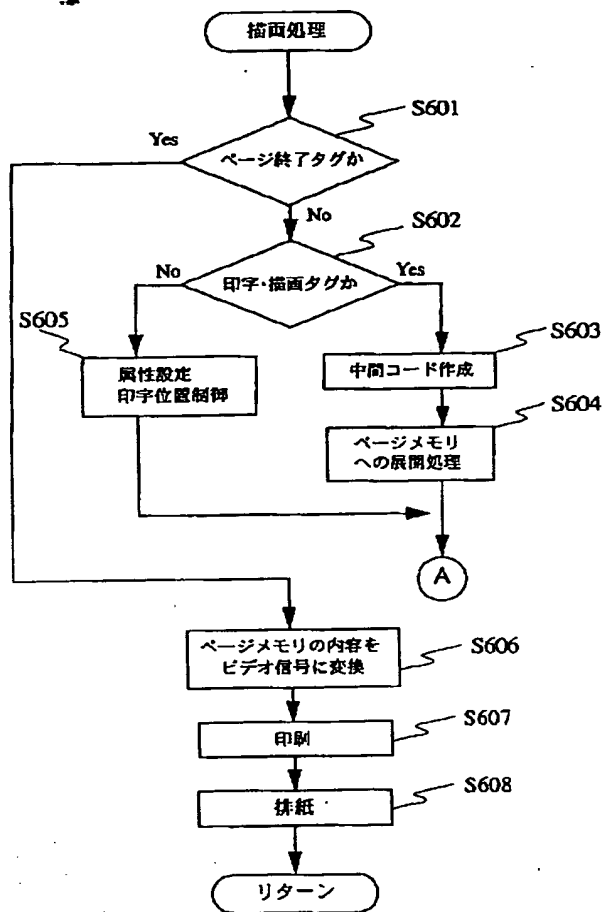
【図4】



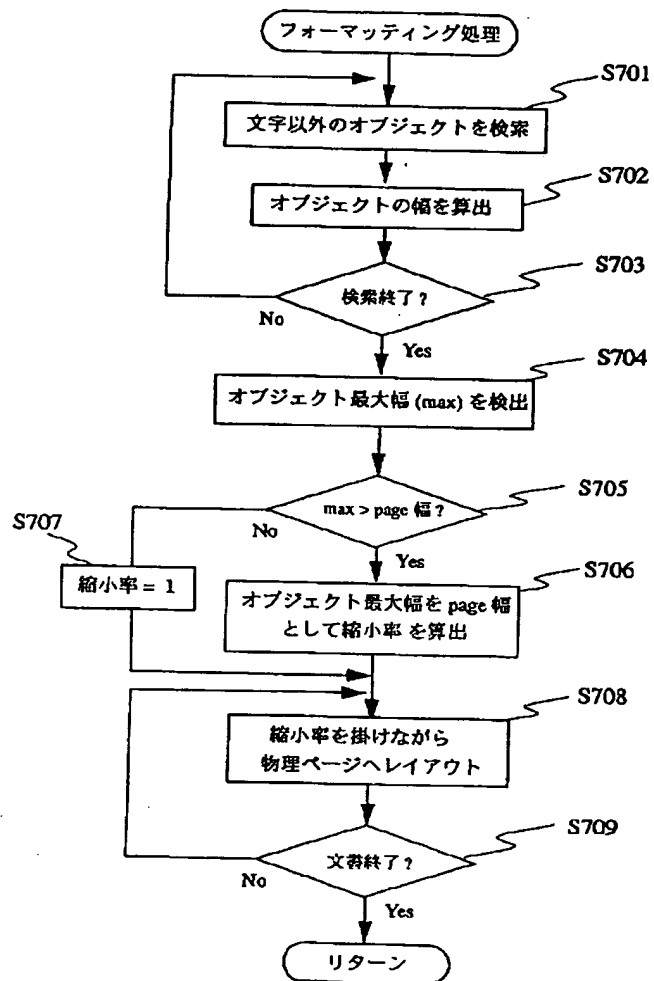
【図5】



【図6】



【図7】



【図11】

```
<html>
<body bgcolor="#ffffff">
<h1>Sample</h1>
This document is written in Markup Language. logical data
structure and logical layout for web browser.
<table border="0" width="100%">
<tr>
<td width="60%"><p align="center">.....</td>
<td width="20%"><p align="right">.....</td>
</tr>
</table>
<hr>
<p><font-size="4"> table 1 </font></p>
</body>
</html>
```

【図12】

```

<document>
<unit size="mm"/>
<text size="28pt" color="red" x="100" y="0">
Sample</text>
<fill pat="aa aa" color="blue"></fill>
<rect 10 200 200 500/>
<text size="12pt" color="black" x="0" y="1">
This document is written </text>
<text size="12pt" x="0" y="2">
in Markup Language. </text>
<text size="12pt" x="0" y="3">
logical data structure and </text>
<text size="12pt" x="0" y="4">
logical layout for web </text>
<text size="12pt" x="0" y="5">
browser. </text>
<fill pat="null"/>
<rect 40 50 120 80/>
<line 80 50 80 80/>
<line 40 60 120 60/>
<line 40 70 120 70/>
<rect 40 90 120 120/>
<line 80 90 80 120/>
<line 40 100 120 100/>
<line 40 110 120 110/>
<text size="8pt" x="3" y="10">
table 1 </text>
</document>

```

【图13】

```
<document>
<unit size="mm"/>
<text size="20pt" color="red" x="100" y="0">
Sample</text>
<fill pat="aa aa" color="blue"></fill>
<rect 30 200 200 500/>
<text size="8pt" color="black" x="0" y="1">
This document is written in </text>
<text size="8pt" x="0" y="2">
Markup Language. logical data</text>
<text size="8pt" x="0" y="3">
structure and logical layout for</text>
<text size="8pt" x="0" y="4">
web browser</text>
<fill pat="null"/>
<rect 30 50 230 110/>
<line 80 50 80 110/>
<line 130 50 130 110/>
<line 180 50 180 110/>
<line 30 70 230 70/>
<line 30 90 230 90/>
<text size="6pt" x="3" y="10">
table 1 </text>
</document>
```

【図14】

```

<document>
<unit size="mm"/>
<text size="23pt" color="red" x="100" y="0">
Sample</text>
<fill pat="aa aa" color="blue"></fill>
<rect 30 200 200 500/>
<text size="10pt" color="black" x="0" y="1">
This document is written in</text>
<text size="10pt" x="0" y="2">
Markup Language,logical </text>
<text size="10pt" x="0" y="3">
data structure and logical </text>
<text size="10pt" x="0" y="4">
layout for web browser</text>
<fill pat="null"/>
<rect 0 50 240 110/>
<line 60 50 60 110/>
<line 120 50 120 110/>
<line 180 50 180 110/>
<line 0 70 240 70/>
<line 0 90 240 90/>
<text size="7pt" x="3" y="10">
table 1 </text>
</document>

```

【図15】

Sample	
This document is written in Markup Language. logical data structure and logical layout for web browser.	
table 1	

【図16】

Sample			
This document is written in Markup Language, logical data structure and logical layout for web browser.			
table 1			

【図17】

Sample			
This document is written in Markup Language. logical data structure and logical layout for web browser.			
table 1			

